

An Empirical Assessment of Colour Use on the WWW

Thomas Mandl

Information Science
University of Hildesheim
Marienburger Platz 22
31141 Hildesheim
Germany
mandl@uni-hildesheim.de

Maximilian Eibl

Social Science Information Center
Schiffbauerdamm 19
10117 Berlin
and
Technical University Chemnitz
Faculty of Computer Science
Straße der Nationen 62
09111 Chemnitz
Germany
eibl@berlin.iz-soz.de

Abstract

Colour is one of the most important design elements for web design and a main determinant for the user experience. Our study analyzes colour use by web designers by mining for pages and extracting the colour information. We included different categories of sites into this study. Colour use statistics and differences in colour use between categories are presented. The consequences for web designers are discussed. Further lines of research at the intersection of human-computer interaction and web mining are outlined. We propose design mining as part of web mining to gain further insight into the design of the web.

1 Colours on the WWW

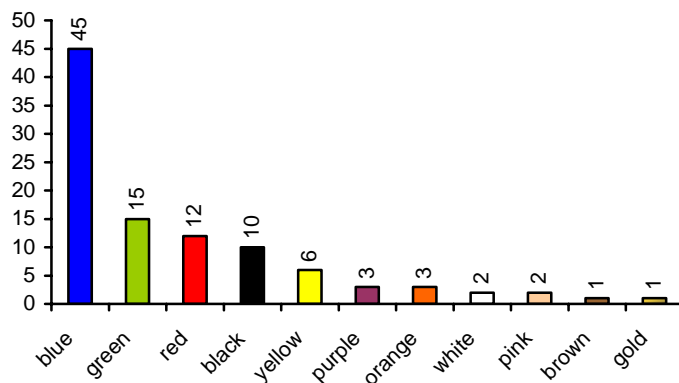
Colour is an important design feature in human-computer interaction and especially in web design (Shubin et al. 1996, Post 1997). Not only can web browsers display colour but virtually any element in HTML can be assigned a colour to it (Niederst 1999).

Experiments have shown that colour and graphics are important for the perception and evaluation of web sites by users (Fogg et al. 2002). The impact of colours seems to be strongly related to the concept of favourite colours. Web sites employing blue as primary colour are

usually seen as favourable compared to those employing alternative colour schemes. Heller 2000 presents a study which asked people about their favourite colour: 45% consider blue to be their favourite colour (see figure 1).

Even minor aspects of colours in web design have attracted research, e.g. the colour of links (Halverson & Hornof 2004). Web style guides provide designers with guidelines on the use of colours on the web. However, little is known about the actual use of colours on the web and the overall distribution of

Figure 1: Favourite colours (Heller 2000: 48)



colours over sites. Empirical research about colours on the web is either limited to small number of pages (Kangas 2001).

Most guidelines for web design mention colour and give some advice to it (e.g. Brinck et al. 2002). Suggestions from style guides are often based on the biological and physical foundations and psychological effects of colours or cultural conventions about the meaning of colours. A biological constraint of the use of colours is for example the ability of humans to distinguish only a rather small set of colours (Regan 2000). A physical effect of colours is the chromostereopsis effect: due to very different wave lengths certain colour combinations cannot be seen sharply at one time: Just think about the trouble reading green text on red paper. The meaning of colours can be a totally different on in two cultures (Romberg et al. 1999). Whereas in western civilization black is always associated with death, Egyptians associate it with resurrection. In India white is associated with death. European household appliances like washing machines are often white. Because household appliances are a typical marriage gift in India, the colour needed to be changed. Otherwise, the marriage gift would signify death for the newly married couple.

2 Evaluation Method

Experiments have shown that colour and graphics are important for the perception and evaluation of web pages by users. This study applies web mining to determine how colours are used in HTML documents. Web mining extracts knowledge from the web which is not explicitly stored by which needs to be distilled from many pages.

2.1 Selected Pages

For this study the web sites referenced in the German version of the Yahoo web catalogue (<http://www.yahoo.de>) for the following five thematic areas were considered:

- Rock & Pop
- Sex
- Programming
- Politics
- University

The purpose of this selection was to analyze sites which highly differ in their characteristics: respectable vs. dubious, work vs. leisure, entertainment vs. education. Do these different site types also differ in the use of colour? If so, further research needs to be done to figure out other typical colour scenarios. That way, page designers can check whether they are within the typical span of colours used in this area or if their colour use is atypical for this thematic domain. In the latter case, they can decide to change their design or to consciously create a new colour experience in order to call the attention of users.

We included pornographic sites because they are an important economic factor on the web and little research is dedicated to them. They are responsible for a very large number of spam pages for content and link spam (Fetterley et al. 2003).

2.2 Applied Software

For this analysis, we adapted software developed for structure and link analysis (Mandl 2003a). This software is based on the open source crawler JoBo (Matuschek 2002). JoBo traverses the web from a given starting point and allows the developer to handle pages based on his needs. It provides document handler which can access a Tidy-Parser to analyze the DOM structure of the HTML pages. From the DOM structure, tag information can be extracted easily.

2.3 Mining Approach

We applied a two-step method. First, the URLs in the Yahoo pages were collected. Then each URL was visited, downloaded and analysed.

3 Results and Discussion

The results show how many pages contain colour coding and how many colours they use. Special focus is given to background and font colours. In most cases, the standard colours are used. However, the cases where non-standard colours are used for background and font tell us a lot about preferred colours.

The results also show differences in colour use between the categories. University sites use white as a background colour most often. Only one third of the sex related sites analyzed use black as font colour whereas some 66% of all other sites use it.

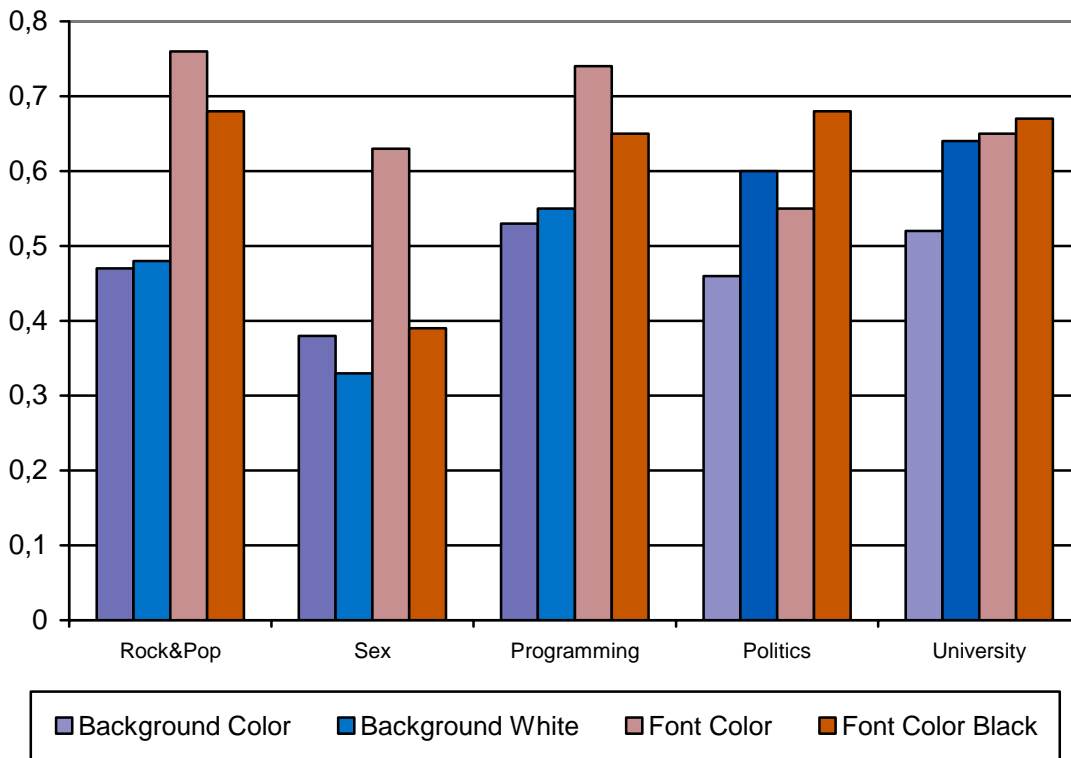


Figure 2: Results of the analysis

The results shown in figure 2 reveal two astonishing facts: One would assume that the use of colours in University sites and sites concerning programming would be nearly equivalent. But the study reveals two other relations: the use of colour in University sites is quite similar to the use of colour in political sites. And sites concerning programming are quite similar to Rock&Pop sites. Especially this last fact seems surprising. Less surprising is the special use of colour in sex related sites – they follow their own rules.

4 Consequences for Web Designers

As the results have shown, web sites highly differ in their use of colour. Colour itself is one of the most important criteria in web design. It highly influences the user's judgment of aesthetics, respectability, credibility and even more: Users even estimate the usability of a web site according to the use of colour - even though there is no rational linkage between usability and the use of colours.

The effects of colours on the judgment of web sites are determined by several factors like the above mentioned biological, cultural or psychological features. Another important factor is user expectation: If the user frequently visits for example University pages, he will build up the expectation that these sites generally use black font colour on a white background because most of the University sites use this colour combination. If a University decides to introduce red as a dominant colour, the user will be at least astonished because he simply did not expect it. If he surfs on sex related sites he will not even notice an unusual colour scheme, just because there is not such scheme. Thus, web designers should know the usual colour schemes – the de facto standards – of the sites they design.

5 Further Lines of Research: Web Design Mining

Empirical research on the web has so far been mainly focused on the user experience for particular web sites or on web usage and log-file analysis. The availability of many pages has led to empirical research on the web in many other areas. For example, the web has been explored as a link network (Barabási 2002) with distinct communities (Flake et al. 2000). Most of this work has been carried out within information retrieval and most metrics developed for the web stem from information retrieval consequently (Dhyani et al. 2002).

5.1 Previous Work on Web Design Mining

More research should be dedicated to empirical web design which analyses the design of many pages by web mining techniques. That way, more knowledge on the design of sites present on the web could be derived. Some examples for studies have been published.

Some studies apply design analysis intellectually. As a consequence, these studies are limited to small numbers of pages.

An interesting experiment tried to correlate the design features to the popularity of the pages (Bucy et al. 1999). Some 5000 pages were manually downloaded and evaluated. Among the features were banners, frames, logos, dominating colour, presence of graphics and the presentation of links. The search for correlations considered the frequency of access to these pages. This metric for popularity was compared to six variables: presence of graphical elements,

dynamic design elements, asynchronous interactive features, real-time interactive features and a combined omnibus structure variable (Bucy et al. 1999).

In order to determine cultural differences in web design between similar cultures, a study on 1000 web sites from Germany and Great Britain was carried out (Hodemacher 2004). It showed interesting results and differences in page layout and site design which could in part be related to the cultural features of the cultures under question.

In an automatic analysis of random pages from the German web, the page structure was the goal of the study (Mandl 2003b). The number of nodes in the document object model (DOM) derived from the page, the number of tables and the number of meta tags were extracted. The two first parameters are indicators for the structural complexity of a page whereas the number of meta tags indicates the extent to which the author aids indexing. An analysis of the pages shows that pages on higher levels in the site structure tree tend to have more DOM elements with a lower deviation. They also seem to have more tables. However, these trends are very weak (Mandl 2003b).

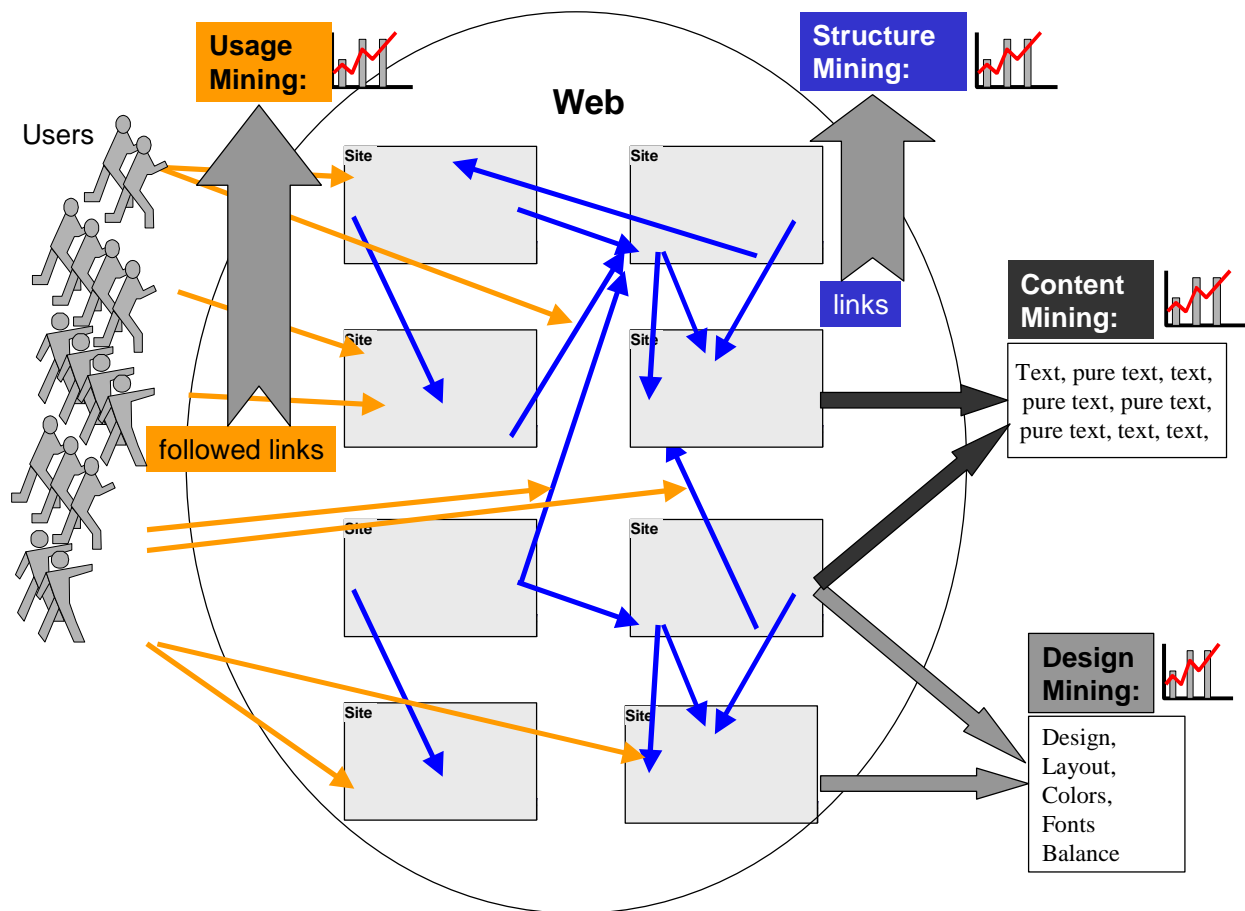


Figure 2: Design Mining in the context of Web Mining

5.2 Further Research on Colours within Web Design Mining

This paper also contributes to empirical web design and demonstrates the use of colour in web sites. The here presented study is a first step of a extensive analysis of the de facto standards in web site colour use. By now the investigated sites are limited to five categories and only black and white font and background colour were taken into account. But though the study was rather limited, web designers who were access given to our results were extraordinary keen on further results. Thus, we are currently planning a comprehensive study including the full set of colours, more categories from the Yahoo catalogue and cross-cultural differences.

The colours coded in the HTML are only one source colour expression of web designers. Many spatial colour effects are brought to the user by graphics. In our continued work, we intend to extract colour histograms (del Bimbo 1999) from graphics and weigh them according to their size and position in the web page. The RGB values can be easily extracted from graphics (Knudsen 1999).

6 Acknowledgement

While involved in this research, the first author was supported by a grant from the German Research Foundation (DFG, grant nr. MA 2411/3-1).

References

- Barabási, Albert-László (2002). *Linked: The New Science of Networks*. Cambridge: Perseus.
- Del Bimbo, Alberto (1999). *Visual Information Retrieval*. Morgan Kaufmann: San Francisco.
- Bucy, Erik P.; Lang, Annie; Potter, Robert F.; Grabe, Maria E. (1999). Formal Features of Cyberspace: Relationships between Web Page Complexity and Site Traffic. *Journal of the American Society for Information Science*, 50, 1246-1256.
- Brinck, Tom; Gergle, Darren; Wood, Scott (2002). *Designing Web Sites that Work: Usability for the Web*. Morgan Kaufmann: San Francisco.
- Dhyani, Devanshu; Ng, Wee; Bhowmick, Sourav (2002). A Survey of Web Metrics. *ACM Computing Surveys*, 34 (4) 469-503.
- Fetterly, Dennis; Manasse, Mark; Najork, Marc; Wiener, Janet (2003): A Large-Scale Study of the Evolution of Web Pages. In *Proc Twelfth International World Wide Web Conference (WWW 2003)*, Budapest. 20.-24. May. pp. 669-678. Retrieved February 15, 2005, from <http://www2003.org/cdrom/papers/refereed/p097/P97%20sources/p97-fetterly.html>
- Flake, G.W.; Lawrence, Steve; Giles, C. Lee (2000): Efficient Identification of Web Communities. In: *Proceedings of the Sixth International Conference on Knowledge Discovery and Data Mining*, Boston. 20.-23. August. pp. 150-160.
- Fogg, B.J.; Marable, Leslie; Stanford, Julianne; Tauber, Ellen R. (2002). *How Do People Evaluate A Web Site's Credibility? Results From A Large Study*. In: *Consumer Web Watch*. Retrieved February 15, 2005, from <http://www.consumerwebwatch.org/dynamic/web-credibility-reports-evaluate-abstract.cfm>
- Heller, Eva (2000). *Wie Farben auf Gefühl und Verstand wirken. Farbpsychologie, Farbsymbolik, Lieblingsfarben, Farbgestaltung*. Droemer.

- Halverson, Tim; Hornof, Anthony J. (2004). Link colours guide a search. In: *Extended abstracts of the 2004 Conference on Human factors and computing systems (ACM SIGCHI)*. Vienna, Austria. pp. 1367 – 1370.
- Hodemacher, Dorte (2004). Der Einfluss von Kultur auf das Webdesign: Eine empirische Untersuchung am Beispiel von Großbritannien und Deutschland. Master Thesis, Information Science, University of Hildesheim.
- Ivory, Melody; Hearst, Marti (2002). *Statistical Profiles of Highly-Rated Sites*. In: Proceedings ACM Conference on Human Factors in Computing Systems (CHI 2002), Mineapolis, USA. 20.-25. April.
- Kangas Steve (2001). *Colour Schemes of Popular Sites*. NetConversions Benchmark Studies #2. Retrieved February 15, 2005, from http://netconversions.com/ip_colourschemes.htm
- Knudsen, Jonathan (1999): JAVA 2D Graphics. Beijing et al.: O'Reilly.
- Mandl, Thomas (2003a). Web Link Behavior and Consequences for Connectivity Based Authority Measures. In *The Twelfth International World Wide Web Conference*. 20-24 May 2003, Budapest. Retrieved February 15, 2005, from <http://www2003.org/cdrom/papers/poster/p204/p204-mandl.html>
- Mandl, Thomas (2003b): Link Analysis and Site Structure in Information Retrieval. In: Dittrich, Klaus; König, Wolfgang; Oberweis, Andreas; Rannenber, Kai; Wahlster, Wolfgang (Eds.), *Informatik 2003: Innovative Informatikanwendungen. Beiträge der 33. Jahrestagung der Gesellschaft für Informatik*. 29. September – 2. October, Frankfurt am Main. Bonn: Köllen pp. 262-267.
- Matuschek, Daniel (2002). JoBo. Retrieved February 15, 2005, from <http://www.matuschek.net/software/jobbo/index.html>
- Niederst, Jennifer (1999). Web Design in a Nutshell. Beijing et al.: O'Reilly.
- Post, David (1997). Colour and Human-Computer Interaction. In Martin Helander, Thomas Landauer, Prasad Prabhu (Eds.), *Handbook of Human-Computer Interaction* (pp. 573-616). Amsterdam et al.: Elsevier.
- Regan, David (2000). Human Perception of Objects: Early Visual Processing of Spatial Form. Sunderland, MA, USA: Sinauer.
- Mandl, T. (2003): Link Analysis and Site Structure in Information Retrieval. In: Dittrich, K.; König, W.; Oberweis, A.; Rannenber, K.; Wahlster, W. (Eds.): *Informatik 2003: Innovative Informatikanwendungen. Beiträge 33. Jahrestagung der Gesellschaft für Informatik* pp. 262-267.
- Romberg M.; Röse, K.; Zühlke, D. (1999). Global Demands of non-european Markets for the Design of User-Interfaces. *MMI-Interaktiv*. Nr. 1, March 1999.
- Shubin, Hal; Falck, Deborah; Johansen, Ati Gropius (1996). Exploring colour in interface design. *interactions* 3. 36 - 48